AD-A058 851 SYSTEMS RESEARCH LABS INC DAYTON OHIO F/G 14/1
VALIDATION OF AN EXPERT ESTIMATE TECHNIQUE FOR PREDICTING MANPO--ETC(U)
MAY 78 D W SAUER, W B ASKREN F33615-76-C-0042

UNCLASSIFIED AFHRL-TR-78-19 NL

1 OF 1
AD
A058851

END
DATE
FILMED
11-78
DDC

AFHRL-TR-78-19

LEVEL II

(2)

AIR FORCE HUMAN RESOURCES LABORATORY

VALIDATION OF AN EXPERT ESTIMATE TECHNIQUE FOR PREDICTING MANPOWER, MAINTENANCE, AND TRAINING REQUIREMENTS FOR PROPOSED AIR FORCE SYSTEMS

By

Daniel W. Sauer
Systems Research Laboratories, Inc.
2800 Indian Ripple Road
Dayton, Ohio 45440

William B. Askren

ADVANCED SYSTEMS DIVISION
Wright-Patterson Air Force Base, Ohio 45433

May 1978
Final Report for Period December 1976 – November 1977

Approved for public release; distribution unlimited.

AIR FORCE SYSTEMS COMMAND
BROOKS AIR FORCE BASE, TEXAS 78235

NOTICE

This final report was submitted by Systems Research Laboratories, Inc., 2800 Indian Ripple Road, Dayton, Ohio 45440, under contract F33615-76-C-0042, project 1124, with Advanced Systems Division, Air Force Human Resources Laboratory (AFSC), Wright-Patterson Air Force Base, Ohio 45433. Dr. William B. Askren, Personnel and Training Requirements Branch, was the contract monitor.

This technical report has been reviewed and is approved for publication.

GORDON A. ECKSTRAND, Director
Advanced Systems Division

DAN D. FULGHAM, Colonel, USAF
Commander

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (*When Data Entered*)

(9) REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER (18) AFHRL TR-78-19	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) (6) VALIDATION OF AN EXPERT ESTIMATE TECHNIQUE FOR PREDICTING MANPOWER, MAINTENANCE, AND TRAINING REQUIREMENTS FOR PROPOSED AIR FORCE SYSTEMS.		5. TYPE OF REPORT & PERIOD COVERED (9) Final rept. Dec[redacted]76 – Nov[redacted]977
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) (10) Daniel W. Bauer, William B. Askren		8. CONTRACT OR GRANT NUMBER(s) (15) F33615-76-C-0042
9. PERFORMING ORGANIZATION NAME AND ADDRESS Systems Research Laboratories, Inc. 2800 Indian Ripple Road Dayton, Ohio 45440		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 62205F (16) 11240107 (17) 01
11. CONTROLLING OFFICE NAME AND ADDRESS HQ Air Force Human Resources Laboratory (AFSC) Brooks Air Force Base, Texas 78235		12. REPORT DATE (11) May [redacted]78
		13. NUMBER OF PAGES 38
14. MONITORING AGENCY NAME & ADDRESS (*if different from Controlling Office*) Advanced Systems Division Air Force Human Resources Laboratory Wright-Patterson Air Force Base, Ohio 45433		15. SECURITY CLASS. (*of this report*) Unclassified (12) 40p.
		15a. DECLASSIFICATION DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (*of this Report*)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (*of the abstract entered in Block 20, if different from Report*)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (*Continue on reverse side if necessary and identify by block number*)

AFSC	maintenance man-hours	subjective estimation
career field	maintenance manpower	task difficulty
crew size	maintenance task time	training facilities/equipment
engineering description package	skill level	training time
expert estimate technique		

20. ABSTRACT (*Continue on reverse side if necessary and identify by block number*)

The objectives were to determine the validity of an expert estimate technique for predicting manpower, maintenance, and training requirements for equipment in the early stages of design, and to develop a guide for implementing the technique. Sixty Air Force technicians from two avionics AFSCs participated as expert estimators. Twenty of these technicians were also qualified avionics instructors. The 60 technicians made estimates of manpower, maintenance, and training requirements using only an engineering description of an operational avionics component. The description contained information available during the early design stages of the component. The accuracy of the estimates was determined by comparing them with manpower, maintenance, and training data associated with the operational equipment. The results indicate that maintenance task time, crew size, skill level,

over

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (*When Data Entered*)

340 400 Gue

Item 20 Continued:

career field, and task difficulty can be estimated with a satisfactory degree of accuracy. Training times were greatly overestimated. The estimates of required training facilities/equipment and the impact of design features on maintenance were nonconclusive. The results also indicate that ten technicians with skill level 5 and with system experience on equipment similar to the proposed equipment will produce acceptable estimates. A prototype guide for using the technique was prepared.

SUMMARY

Objective

The objectives were to determine the validity of an expert estimate technique for predicting the manpower, maintenance, and training requirements for equipment which is in the early phases of design, and to develop a guide for implementing the technique.

Approach

Experienced Air Force maintenance technicians estimated the manpower, maintenance, and training requirements for aircraft avionics equipment using only an early design phase description of that equipment as the basis for the estimates. The accuracy of their estimates was determined by comparing manpower, maintenance, and training data associated with operational use of the equipment with the estimates.

An engineering description, containing only information which had been available during the early design stages, was developed for the AN/APN-147(V) Doppler Radar Navigation System. This description served as the stimulus material for the estimates. A questionnaire was designed to collect estimates of: maintenance man-hours; crew size; skill level; career field; task difficulty; training time for resident, on-the-job training (OJT), and field training detachment (FTD) training; and training facilities/equipment. Other sections of the questionnaire asked technicians to evaluate the features of the design itself in terms of impact on maintenance time, task difficulty, errors, and safety hazards.

Sixty technicians from two career fields, the Avionic Navigation Systems Specialist (328X1) and the Avionic Inertial and Radar Navigation System Specialist (328X4), with varying levels of experience in avionics maintenance and training, participated as the expert estimators. The technicians were located at four Air Force bases.

Statistical comparisons were made between the estimates and the operational criteria data in order to determine the accuracy of the estimates. Analyses of the data were also made to determine the effect on accuracy of the factors of kind and amount of weapon system experience of the estimators, and quantity of estimators. A guide was developed based on the procedures followed, the experiences gained, and the results of the study.

Background

The human resources associated with support of Air Force systems can account for much of the system readiness and cost. As systems become more complex and more expensive, it becomes necessary to optimize

the use of human resources. Knowing in advance how a proposed system could affect the human resources, the designers and engineers could change, or modify, those aspects of the system which have an adverse effect. The earlier in the system design these impacts are known, the easier and less costly it is to make the required changes. Subjective estimates, or expert estimates appear to be an inexpensive and fast method to predict the human resource requirements for proposed system equipment. The question is, how accurate are these estimates? This study was conducted to expand the relatively small body of evidence on the validity of the expert estimate technique, and to develop and document a standard technique for generating the estimates.

Results and Conclusions

Comparisons of the estimates with the **criteria** data indicate that maintenance man-hours, crew size, skill level, career field, and task difficulty can be estimated with a satisfactory degree of accuracy. However, the technicians greatly overestimated the training time for resident, field training detachment training, and on-the-job training. Estimates of the training facilities/equipment and the impact of design features on maintenance were inconclusive.

No differences in accuracy of estimates were found between three groups of technicians with different kinds of weapon system experience. No differences in accuracy were found between technicians with four or fewer years of systems experience and technicians with over four years of experience. A minimum of 10 estimators is recommended to produce stable estimates. It is recommended that estimators be at least skill level 5 and have maintenance experience on systems or equipment similar to the proposed equipment.

A prototype guide for implementing the expert estimate method was developed and is available as AFHRL-TR-78-19 (Supplement 1), from the Defense Documentation Center and NTIS. The guide includes the five topics of: (1) developing the engineering description package of the proposed design; (2) developing the questionnaire to collect the estimate data; (3) selecting the technicians to serve as expert estimators; (4) collecting the estimate data; and (5) analyzing the data and preparing it for use by the engineer. The guide also includes a copy of the engineering description package and the data forms used by the expert estimators.

PREFACE

This study was performed by Systems Research Laboratories, Inc. (SRL), 2800 Indian Ripple Road, Dayton, Ohio. Technical direction was provided by the Advanced Systems Division, Air Force Human Resources Laboratory (AFHRL), Wright-Patterson Air Force Base, Ohio.

The AFHRL support was provided under project 1124, Human Resources in Aerospace System Development and Operations, Dr. Ross L. Morgan, Project Scientist; and task 112401, Personnel, Training, and Manning Factors in the Conception and Development of Aerospace Systems, Dr. William B. Askren, Task Scientist.

Systems Research Laboratories, Inc. performed the research under contract F33615-76-C-0042 with Dr. Norman R. Potter as Principal Investigator and Mr. Daniel W. Sauer as Associate Principal Investigator.

The authors wish to acknowledge the many individuals who contributed to this effort. Appreciation is extended to the Strategic Air Command personnel in the 301st AMS at Rickenbacker AFB, Ohio and in the 305th AMS at Grissom AFB, Indiana. Appreciation is also extended to Military Airlift Command personnel from the 438 AMS at McGuire AFB, New Jersey, to training personnel from the 3300 Technical Training Wing at Keesler AFB, Mississippi, and to avionics maintenance personnel from the Springfield, Ohio, Air National Guard Unit. Special appreciation is extended to the Canadian Marconi Company for their guidance in preparation of the Engineering Description Package.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF TABLES (cont)

VALIDATION OF AN EXPERT ESTIMATE TECHNIQUE FOR PREDICTING MANPOWER, MAINTENANCE AND TRAINING REQUIREMENTS FOR PROPOSED AIR FORCE SYSTEMS

INTRODUCTION

Background and Problems

One of the biggest challenges facing Air Force system designers and engineers is that of obtaining optimum performance from new systems while keeping costs within strict budget limitations. The human resources associated with Air Force systems can contribute to both system performance and system cost and as such, represent an important area for affecting performance increases and cost savings.

If human resources do have such a large impact on system costs, then it follows that this area should be examined carefully to ensure optimum utilization of human resources and to determine potential cost savings. In 1963, Haines and Gael stated that "There is an urgent need for better use of our manpower," and advocated that a "simple and accurate" technique was needed for obtaining estimates of manpower requirements for new systems early in the design phase of these systems. They maintained that trade-offs between design requirements and manning requirements could be made more effectively and economically in these early design stages. They surveyed industry to determine which of five human resource estimation techniques were most frequently employed. The techniques included expert estimation, historical comparison, task analysis, sovereign factor and models. The technique which appeared to be the simplest, expert estimation, was one of the least frequently employed techniques in industry. Although expert estimation seems to be a simple technique, it is probable that industry was reluctant to employ the technique because so little was known concerning its accuracy.

Smith, Blanchard, and Westland (1971) recognized the potential value of expert estimate techniques for providing data for man-machine models. They studied restore time estimates made by technicians for a variety of radar, radio, and computer equipment. The technicians used the expert estimate approach. Although Smith et al. were not able to find satisfactory data with which to validate the technicians' estimates, they cautiously concluded that the technique appeared to be a "highly successful, cost-effective approach to obtaining repair time data."

More recently, Whalen and Askren (1974) contributed to the data on expert estimation techniques in several areas. They obtained time estimates for a variety of maintenance tasks as well as estimates of crew size, skill level, and career field for selected Air Force systems. Furthermore, the estimates were based on conceptual phase engineering descriptions of the equipment thereby providing data on expert estimates

made well before actual production had begun. Finally, they were able to find suitable data for validating the expert estimates. They reported that, overall, technicians tended to underestimate maintenance times by 29.4 percent and produced highly accurate estimates of crew size, skill level, and Air Force Specialty Code (AFSC).

The concept of a simple and accurate method for estimating human resource requirements is still very appealing and it appears that the expert estimate technique has the potential to become that simple and accurate method. Yet, the evidence to date has not been sufficient to recommend the expert estimate technique for applications during development of new weapon systems.

Objectives

The objectives of this study were to provide additional evidence regarding the validity of expert estimates of human resources data and to refine and extend an expert estimate technique which was developed by Whalen and Askren (1974). Specifically, additional validity data were to be collected for man-hours, skill level, crew size and AFSC estimates. The Whalen and Askren (1974) technique was to be refined and extended to include estimates of training requirements and to determine the qualities and quantities of raters needed to provide useful estimates.

RESEARCH APPROACH

The research approach used was, in most respects, similar to a longitudinal study. A true longitudinal study would have technicians estimate the manpower, maintenance, and training data for a system in the early design stages and, some years hence, when the system became operational, would have compared the estimated data with the actual data to determine the validity of the estimates. Available time, however, precluded a true longitudinal study. The research approach used could be called a reverse or backward longitudinal study.

The term reverse or backward longitudinal study comes from the fact that an engineering description of a piece of operational equipment was prepared to include only information that would have been available during the early design phase of that equipment. Using this specially prepared engineering description only, technicians with various amounts and kinds of work experience estimated the manpower, maintenance, and training requirements for that equipment. The technicians' estimates were then compared to manpower, maintenance, and training data available on the operational equipment to determine the validity of those estimates.

Selection of Operational Equipment

The frequency tracker in the AN/APN-147(V) Doppler Radar Navigation System was chosen as the test bed equipment. The choice of a piece of avionics equipment was made since it has been shown that the avionics technology area has the greatest impact on human resources data (Whalen and Askren, 1974).

Development of Engineering Description Package

The engineering description package was modeled after a similar description package used by Whalen and Askren (1974). The Canadian Marconi Company, manufacturer of the AN/APN-147(V) radar set, was consulted to determine the type of information which appeared in early design stage descriptions of the radar set. This information was used by Systems Research Laboratories, Inc. (SRL) avionics engineers to construct the engineering description package for the frequency tracker. The engineering description package contained a section on theory of operation, physical and functional descriptions of the components, illustrations, circuit diagrams, descriptions of test equipment, built-in test features, and information on maintainability.

Estimating Manpower, Maintenance, and Training Data Items

The manpower, maintenance, and training data items for which estimates were collected are presented in Table 1. Many of the manpower

TABLE 1. MANPOWER, MAINTENANCE, AND TRAINING ITEMS

Maintenance Man-Hours
Crew Size
Skill Level
Career Field
Task Difficulty
Training Time by:
Training Type
Training Content
Training Facilities/Equipment
Design Feature Impact on:
Maintenance Difficulty
Safety Hazards
Maintenance Time
Maintenance Errors

data items have been included in previous research on expert estimation techniques (Smith, Blanchard, and Westland, 1971; Whalen and Askren, 1974). Most of the training items and the maintenance items have not appeared in previous expert estimate research. They were included in this study to determine the accuracy with which technicians could estimate these types of data.

One of the objectives of this research effort was to determine what scaling techniques or decision-making aids could be used to facilitate the collection of manpower, maintenance, and training data estimates. Smith, Blanchard, and Westland (1971) considered estimates of restore time to be made by a "method of direct quantification" or direct estimates. They collected estimates of equipment restore time for the 1st, 25th, 50th, 75th, and 99th percentiles. They found, however, that technicians had trouble understanding the concept of percentiles within the relatively short time available to present the concept. For this study, a similar but less complex approach was used. Technicians were asked to estimate the average repair times as well as the shortest and longest repair times for a given task. Although the focus of this study was on measures of central tendencies, it was felt that having technicians consider both extremes of the repair time range may serve to improve their estimates of the average repair times.

Since training times are not ususally subject to the variations possible in repair times, the above approach was limited to estimates of repair times. Training times were estimated directly with no attempt to estimate a distribution of training times.

Maintenance task difficulty was represented by a continuous scale 100 mm long. Huber and Delbecq (1972) have shown that accuracy of group judgments improved when a continuous rather than a discrete scale

is used. The words "Very Easy" (0 mm mark), "Average Difficulty" (50 mm mark), and "Very Difficult" (100 mm mark) were added as anchors to indicate the direction of the difficulty scale.

It was recognized that crew sizes may vary for any given task depending on the severity of the problem, available personnel or individual squadron policy. Based on discussions of using personal probabilities in decision-making research (Decisions and Designs, Inc., 1973), it appeared that using the probability approach would be useful in estimating crew sizes for the defined maintenance tasks. In estimating the probability of a one-, two-, three- and four-man crew, it seemed that technicians could produce a more meaningful and useful estimate than if only a single estimate of crew size were made.

Estimates of skill level, career field, required training facilities and equipment, and the impact of design features were considered items which would be directly estimated without employing additional techniques. In the cases of design feature evaluation and training equipment requirements, an open-ended response format was used to allow for the wide range of potential responses.

The manpower, maintenance, and training item questions were combined in a questionnaire booklet. A page for collecting background data on each technician was also included in the questionnaire. Additional pages were included to allow estimators to evaluate the engineering description package.

The following paragraphs define the manpower, maintenance, and training items that were estimated. They also describe the data used to validate these estimates.

Maintenance Man-Hour Estimates and Criteria. Maintenance man-hours were defined as the number of man-hours necessary to complete a specified maintenance task. To arrive at man-hour estimates, technicians were first requested to estimate the actual task times for six maintenance tasks involving the APN-147 doppler radar (see Table 2). Man-hours for each of the tasks were calculated by multiplying the maintenance task time estimate by the crew size estimate as described below.

The maintenance man-hour criteria data for the six tasks (Table 2) were taken from AFM 66-1 maintenance data for the frequency tracker from the AN/APN-147(V) radar set. The tasks were selected to represent a variety of flight line and shop maintenance activities. The maintenance man-hours for the selected tasks represented maintenance actions performed in the 12-month period from 1 May 1976 to 30 April 1977.

Crew Size Estimates and Criteria. Crew size was defined as the number of technicians required to perform the specified maintenance

TABLE 2. DESCRIPTIONS OF MAINTENANCE TASKS FOR WHICH TECHNICIANS ESTIMATED MAN-HOURS, CREW SIZE, SKILL LEVEL, AND TASK DIFFICULTY

Task Characteristics	Maintenance Tasks 1	2	3	4	5	6
Equipment Name	Doppler Signal Processor	Doppler Signal Processor	Doppler Signal Processor	Doppler Signal Processor	Power Supply Module	Signal Comparator Module
Type Malfunction	No Output	Lock On Malfunction	Lock On Malfunction	Lock On Malfunction	No Output	Lock On Malfunction
Maintenance Location	Flight Line	Flight Line	Flight Line	Shop	Shop	Shop
Maintenance Action Taken	Remove and Replace	Remove	Adjust	Repair	Bench Check-Repair Deferred	Repair

task. Technicians estimated the probability that a one-, two-, three-, and four-man crew would be required to perform the specified maintenance task. Criteria data on crew size were collected from four maintenance personnel highly experienced with the AN/APN-147(V) radar set. These maintenance personnel became the Manpower and Maintenance criteria data judges (Table 3). These judges, as a group, established the probabilities associated with the four crew sizes based on extensive squadron experience with the AN/APN-147(V) radar set.

Skill Level Estimates and Criteria. Skill level was defined in terms of the Air Force skill level classifications of 3, 5, 7, and 9. Technicians made direct skill level estimates in conjunction with their crew size estimates for the six maintenance tasks (Table 2). The Manpower and Maintenance judges (Table 3), as a group, established the skill level criteria data for each maintenance task.

Career Field Estimates and Criteria. Career field was defined as the five-digit AFSC. The estimates of career field were made in two ways. First, technicians were asked to indicate what type of changes, if any, would be required within the family of 328XX AFSCs (avionics and electronics) to maintain the proposed equipment. The second approach was to have technicians directly estimate which one of five AFSCs from the family of 328XX AFSCs was most appropriate for work on the equipment. The Manpower and Maintenance judges identified the specific AFSC held by technicians actually performing maintenance on the AN/APN-147 radar set.

Task Difficulty Estimates and Criteria. Task difficulty was defined as the degree of difficulty of performing each of the six APN-147 maintenance tasks (Table 2). The technicians rated difficulty on a 100 point rating scale. The Manpower and Maintenance judges, as a group, used the same procedures to establish the task difficulty criteria.

Training Time Estimates and Criteria. Training time estimates were defined as the amount of time individuals would need to spend on various training topics within various types of training situations to become qualified on the proposed equipment at skill level 5. Specifically, technicians were to estimate the number of hours of resident training, on-the-job training (OJT) and field training detachment (FTD) training necessary to qualify an airman just graduated from basic training to 5-level proficiency on the proposed equipment. Time estimates for each training situation were broken down into four general training topics: basic electronics, equipment orientation, theory of operation, and maintenance (Table 4). The general topic of maintenance training was further broken down into ten specific maintenance activities; remove, replace/install, bench check, repair, calibrate, adjust, test/inspect/service, check, assemble, and disassemble (see Table 5).

TABLE 3. CHARACTERISTICS OF CRITERIA DATA JUDGES FOR MANPOWER, MAINTENANCE, AND TRAINING DATA

Criteria Judges' Background	n	Mean Age	Mean Years of Experience	Type of Criteria Provided*
Avionics Maintenance	4	34	9.6	Crew size, skill level, career field, FTD training time, OJT training time, task difficulty, design feature evaluation
Avionics Instructor	4	42	10.2	Resident training time, facilities/equipment

*Remaining criteria, man-hours, obtained from Air Force 66-1 maintenance records.

TABLE 4. FORMAT FOR ESTIMATING TRAINING TIMES: GENERAL TRAINING TOPICS

Training Topics	Types of Training: Resident	OJT	FTD
Basic Electronics	*		
Equipment Orientation			
Theory of Operation			
Maintenance			

*Training time estimates in hours

TABLE 5. FORMAT FOR ESTIMATING TRAINING TIMES
MAINTENANCE TRAINING TOPICS

Maintenance Training Topics	Resident	OJT	FTD
Remove	*		
Replace/Install			
Bench Check			
Repair			
Calibrate			
Adjust			
Test/Inspect/Service			
Check			
Assemble			
Disassemble			

*Training time estimates in hours

Experienced avionics instructors at Keesler Technical Training Center provided the actual resident training time (Tables 4 and 5) for APN-147 technicians. This group of training criteria judges is described in Table 3. Actual training times for OJT and FTD training (Tables 4 and 5) were provided by the manpower and maintenance criteria data judges at McGuire AFB, New Jersey.

Training Facilities/Equipment Estimates and Criteria. Training facilities/equipment were defined as the types of training devices, training equipment and training aids required to support the proposed equipment. Technicians used open-end response formats to make their estimates of the training devices, training aids, and training equipment required to support resident, OJT and FTD training on the proposed equipment. Criteria for validating the training facilities/equipment estimates were provided by avionics instructors at Keesler AFB, Mississippi (see Table 3).

Design Feature Evaluation and Criteria. Design features were defined as any characteristics of the equipment or its components which could be identified as having some effect on the maintenance of that equipment. Design feature evaluations were solicited for any design features of the proposed equipment which, in the opinion of the technicians, would: (1) make maintenance particularly difficult; (2) present safety hazards; (3) unduly increase maintenance time; and (4) contribute to maintenance errors. Technicians used an open-end response format to identify these equipment design features which they felt adversely affected maintenance and maintenance safety. The Manpower and Maintenance judges provided criteria

in the form of lists of design deficiencies affecting the four categories described above.

Subjects

Since an avionics component of an avionics system was the test bed equipment, it was assumed that estimators of the manpower, maintenance, and training data items would come from career fields (AFSCs) and systems most closely related to the test bed equipment. Three groups of avionics technicians were selected to serve as estimators of the manpower, maintenance and training data items: Group D (n = 8) was made up of avionic and inertial navigation system specialists (328X4) with operational maintenance experience on doppler systems other than the APN-147; Group X (n = 35) was made up of avionic navigation systems specialists (328X1) with no doppler maintenance experience; and Group A (n = 17) was made up of avionic and inertial navigation systems specialists (328X4) with operational maintenance experience on the APN-147 doppler radar. These three groups were created to investigate the effects on estimates of manpower and maintenance data items of three levels of systems experience.

For estimating training data items, a fourth group, Group K (n = 20) was formed. Technicians in Group K were resident avionics instructors at Keesler Technical Training Center, Keesler AFB, Mississippi. The instructors in Group K also made estimates of the manpower and maintenance data items and were included in either Groups A, D, or X depending upon their previous operational maintenance experience.

Tables 6 through 8 summarize the qualifications and characteristics of the technicians.

Data Collection

Prior to conducting the actual data collection, the engineering description package, the questionnaire and the collection procedures were tested and refined on a group of avionics technicians from the Ohio Air National Guard at Springfield, Ohio. Having experienced no major problems with either the test instruments or the procedures, the actual data collection effort began.

Visits were made to squadrons at Rickenbacker AFB, Ohio, (301 AMS) Grissom AFB, Indiana (305 AMS), McGuire AFB, New Jersey (438 AMS), and to the 3300 Technical Training Wing at Keesler AFB, Mississippi.

TABLE 6. EXPERT ESTIMATORS: EXPERIENCE AND CAREER FIELDS

Group Code and Experience	Career Field	Number of Estimators for Each Category of Data: Manpower and Maintenance	Training
Group D (Other Doppler)	328X4	8	6
Group X (No Doppler)	328X1	35	20
Group A (APN-147)	328X4	17	14
Group K (Avionics Inst.)	328X1 and 328X4	0	20
	Total*	60	60

*A total of 60 estimators was used. The Group K avionics instructors were also used in Groups A, D, and X depending upon individual maintenance experience to estimate the manpower and maintenance data.

TABLE 7. EXPERT ESTIMATORS: SKILL LEVEL BY SYSTEMS EXPERIENCE GROUPS

	Experience Groups			
Skill Level	APN-147	Other Doppler	No Doppler	Avionics Inst.
5	14	5	15	6
7	2	2	20	14
9	1	1	--	--
Total	17	8	35	20
Mean Skill Level	5.59	6.00	6.06	6.40

TABLE 8. MAINTENANCE EXPERIENCE OF EXPERT ESTIMATORS BY SYSTEMS EXPERIENCE GROUPS

	Avionics Experience							
	APN-147		Other Doppler		No Doppler		Instructors	
	Mean	Range	Mean	Range	Mean	Range	Mean	Range
Years of Maintenance Experience	6.53	1-20	7.25	1-22	9.26	1-24	11.8	2-24

The actual data collection occurred in group settings of from two to ten technicians. Each technician was given a copy of: the Engineering Description Package for the Doppler Signal Processor; the Manpower, Maintenance, and Training Estimate Questionnaire; and the Engineering Description Package Evaluation. The questionnaire administrator briefed the technicians on the purpose of the reseach project prior to giving the verbal instructions for the questionnaire and evaluation. Technicians were advised to read the entire Engineering Description Package before attempting to make their estimates. They were also encouraged to refer to the Engineering Description Package whenever necessary during the session. Upon completion of the questionnaire, the technicians were asked to evaluate the Engineering Description Package for the Doppler Signal Processor.

RESULTS

Maintenance Man-Hour Estimates

The validity of the maintenance man-hour estimates was determined by comparing the estimated man-hours with the actual man-hours reported for each of the six maintenance tasks (Table 2). Accuracy values were calculated by dividing the estimated time by the actual time. An estimate that perfectly predicted man-hours had an accuracy value of 1.00. Estimates below the actual time had accuracy values <1.00 while estimates above the actual time had accuracy values >1.00. Table 9 presents the accuracy values for three groups of technicians. In terms of accuracy, the groups tended to under estimate maintenance man-hours. Group A, with APN-147 experience, had the most accurate time estimates. The group with no doppler experience produced the next best record while the group with other doppler experience produced the worst accuracy record. A Kruskal-Wallis one-way analysis of variance, however, produced an H equal to 2.467 ($p<.30$) indicating that there were no significant differences among the three groups in terms of accuracy of man-hour estimates. It is interesting to note that the accuracy values were close to the accuracy value of .70 found by Whalen and Askren (1974).

TABLE 9. ACCURACY VALUES FOR ESTIMATES OF MAINTENANCE MAN-HOURS BY SYSTEMS EXPERIENCE GROUPS

	Accuracy Values for Man-Hours Estimates						
	Task 1	Task 2	Task 3	Task 4	Task 5	Task 6	Mean Mean
Group A (APN-147	1.05	1.28	.95	.48	.49	.55	.80
Group D (Other Doppler)	.76	.76	.59	.32	.23	.34	.50
Group X (No Doppler)	.82	.86	1.02	.43	.31	.58	.67
Overall	.88	.97	.94	.43	.35	.54	

$$\text{Accuracy Values} = \frac{\text{Estimated Man-Hours}}{\text{Actual Man-Hours}}$$

The interrater reliabilities for single raters ranged between .25 (Group D) and .27 (Groups A and X) for the three groups. The interrater reliability coefficients for the groups using the Spearman-Brown Prophecy formula were: .87 for Group A (APN-147, n = 17), .73 for Group D (other doppler, n = 8), and .97 for Group X (no doppler, n = 35).

An examination of the overall accuracy scores for each task suggests that there may be differences in the accuracy of man-hour estimates based on the type of task for which an estimate is made. Although the small number of group estimates per task did not permit a statistical analysis of these differences, it appears that the flightline maintenance tasks (1, 2, 3) were estimated more accurately than the shop maintenance tasks (4, 5, 6).

The accuracy values for Groups A and X were divided into two maintenance experience groups: technicians who reported four or fewer years of maintenance experience and technicians who reported five or more years of maintenance experience. Since Group D had only two technicians with five or more years of maintenance experience, an analysis was not conducted within this group. The results of the comparisons are presented in Table 10. The Mann-Whitney U Test indicated that there were no significant differences in accuracy values based on length of maintenance experience for either Group A ($U = 7$, $p<.155$) or Group X ($U = 18$, $p<.531$).

TABLE 10. ACCURACY VALUES[1] FOR MAINTENANCE MAN-HOUR ESTIMATES BY LENGTH OF MAINTENANCE EXPERIENCE

Years of Maintenance Experience[2]	Accuracy Values For Man-Hours Estimates						
	Task 1	Task 2	Task 3	Task 4	Task 5	Task 6	Mean
Group A ≤4 yrs	1.25	1.28	–[3]	.59	.61	.63	.70
>4 yrs	.91	1.28		.41	.40	.50	.87
Group X ≤4 yrs	.86	.66	.86	.53	.36	.56	.69
>4 yrs	.81	.98	1.15	.36	.33	.56	.70

[1]Accuracy Value = $\frac{\text{Estimated Man-Hours}}{\text{Actual Man-Hours}}$

[2]Group D not included due to small number of raters in over 4 yrs group.

[3]Data was not usable due to misinterpretation of item.

The effect of group size on accuracy values was also investigated. Huber and Delbecq (1972) indicated that the optimum group size for making accurate decisions was between five and ten. Although they did not have validity data, Smith, Blanchard, and Westland (1971)

felt that groups with 18 to 20 members would produce adequate task time estimates.

To test these findings, accuracy values for groups of 5, 10, and 20 estimators were calculated during this study. The individual estimator's accuracy values from each of the three systems experience groups (A, D, and X) were used as data pools to construct the various sized groups.

Ten groups of 5, 10, and 20 were randomly selected from the time estimate data of Group X (n = 35). Likewise, ten groups of five were randomly selected from Group D (n = 8) and ten groups of five and ten were selected from Group A (n = 17). The mean accuracy values for these groups of 5, 10, or 20 estimators are shown in Table 11. The mean accuracy values for the original groups (A, D, and X) are also presented for reference. These results show that mean accuracy values do not improve for groups larger than five estimators.

TABLE 11. MEAN ACCURACY VALUES FOR ESTIMATES OF MAINTENANCE MAN-HOURS FOR VARYING GROUP SIZES WITHIN SYSTEM EXPERIENCE GROUPS

		Man-Hour Accuracy Values					
	Number of	Group A		Group D		Group K	
Group Size	Groups	Mean	Range	Mean	Range	Mean	Range
5	10	.82	.61-91	.52	.38-.59	.71	.42-1.12
8*	1	-	-	.50	-	-	-
10	10	.74	.63-.84			.67	.54-.80
17**	1	.80	-				
20	10					.70	.55-.80
35***	1					.67	-

*Total n for Group D
**Total n for Group A
***Total n for Group X

However, the range of accuracy values presented for the groups of 5, 10, and 20 estimators (Group X) indicates that as the size of the estimator group is increased, there is less variability in the accuracy values and, therefore, less variability in the man-hour estimates. This seems to hold for group sizes up to 10. Table 12 contains the interrater reliability coefficients, as well as the mean accuracy values, for the various sized estimator groups formed by the Group X technicians. Using the Spearman-Brown Prophecy formula and the interrater reliability coefficient of .27 for an individual Group X estimator, the coefficients increase from a value of .65 for a group of 5 to a value of .93 for a group of 35.

TABLE 12. MEAN ACCURACY[1] AND INTERRATER RELIABILITY COEFFICIENTS FOR VARIOUS SIZED ESTIMATOR GROUPS

Group Size	Number of Groups	Mean Accuracy	Interrater Reliability Coefficient
5	10	.71	.65
10	10	.67	.79
20	10	.70	.88
35	1	.67	.93

[1] $\text{Accuracy Value} = \dfrac{\text{Estimated Man-Hours}}{\text{Actual Man-Hours}}$

At this point it appears that an estimator group size of 10 will give accurate and stable estimates for the practical systems analysis application. Given an abundance of time and resources, estimator groups greater than 10 would add to the confidence in the estimated data.

Crew Size, Skill Level and AFSC Estimates

Crew Size. The crew size estimates represented probability estimates by the technicians. They estimated the probability that one-, two-, three-, and four-man crews would be required to perform each of the six maintenance tasks. These estimates were compared with the criteria probabilities obtained from the manpower and maintenance judges. Table 13 contains the accuracy values for the probability estimates of one- and two-man crews for the three groups of estimators. The criteria probabilities for three- and four-man crews for all tasks were zero. Accuracy values, therefore, could not be calculated for three- and four-man crew estimates. However, the estimated probabilities for crews of three and four ranged from .06 to .00 indicating close agreement with the criteria probabilities.

In general, the estimated probabilities were sufficiently close to the criteria probabilities to warrant use of estimated data in early design studies.

TABLE 13. ACCURACY VALUES[1] FOR ESTIMATES OF CREW SIZES FOR TASKS 1 THROUGH 6

	Task 1 Crew		Task 2 Crew		Task 3 Crew		Task 4 Crew		Task 5 Crew		Task 6 Crew		Overall Crew	
	1	2	1	2	1	2	1	2	1	2	1	2	1	2
Group A (n = 17) (APN-147)	.85	1.47	.84	1.33	1.10	.83	.84	*	1.28	.34	1.25	.42	1.03	.88
Group D (n = 8) (Other Doppler)	.82	1.23	.96	1.29	.83	1.08	.79	*	1.14	.53	1.11	.58	.94	.94
Group X (n = 35) (No Doppler)	.73	1.45	.79	1.36	.49	1.64	.70	*	1.10	.60	.95	.93	.79	1.17
Overall	.78	1.37	.83	1.34	.71	1.34	.75	*	1.16	.56	1.06	.74	.88	1.06

*The probability of a 2-man crew for Task 4 was zero. An accuracy value could not be calculated.

[1]Accuracy Value = $\frac{\text{Estimated Probability of Crew Size}}{\text{Criterion Probability of Crew Size}}$

A Kruskal-Wallis one-way analysis of variance was performed for the accuracy values for crew sizes of one and two. The analysis indicated that there were no significant accuracy differences among the three groups (A, D, and X) for estimates of crew size one ($H = 4.90$, $p<.10$) or for estimates of crew size two ($H = 2.16$, $p<.50$).

Inspection of the accuracy values by maintenance task in Table 13 suggests that the type of task may influence the accuracy of the estimate. Although there was not a sufficient quantity of estimates per task to statistically evaluate the differences, it appears that technicians may be overestimating the probability of a crew size of two for Tasks 1 through 3 while underestimating the probability of a crew of one for these tasks. The opposite appears to hold for Tasks 5 and 6.

Tasks 1 through 3 are flightline tasks while Tasks 5 and 6 are performed in the maintenance shop. These task differences may be a factor contributing to the differences in the estimates.

Skill Level. Skill level estimates were linked to crew size estimates in that when a technician estimated the probabilities of a one-, two-, three or four-man crew for each of the tasks, he also estimated the skill levels required for that one-, two-, three- or four-man crew. These estimates were compared with the criteria skill levels. The percentage of correct estimates served as the indicator of accuracy. Since the criteria indicated that no three- or four-man crews would be required to perform the tasks the skill level estimates were analyzed for one- and two-man crews only.

Table 14 contains these percentages of correct skill level estimates for one- and two-man crews for the six maintenance tasks. For crew size one, the mean percentages of correct estimates is very high, 88 to 100 percent, for all three groups. For crew size two, the mean percentages of correct estimates falls between 59 percent and 68 percent for the groups.

An inspection of skill level estimates for crew size of one indicates essentially no differences in terms of correct estimates among the three estimator groups. A Kruskal-Wallis analysis of the crew two skill level estimates also indicates no differences among the three estimator groups ($H = .32$, $p<.90$).

Career Field. Estimates of AFSC were made at two levels of specificity. First, technicians were told that the maintenance personnel for the proposed system would come from the 328XX career field (avionics and electronics). The technicians were to estimate whether any changes to the AFSCs in this career field would be required to support the proposed equipment. The responses for the three systems experience groups are contained in Table 15. The responses to this item are difficult to interpret in terms of accuracy. Since the

TABLE 14. ACCURACY[1] OF SKILL LEVEL ESTIMATES

Task	Criteria Skill Levels	Estimator Groups: Group A (n = 17) (APN-147)	Group D (n = 8) (Other Doppler)	Group X (n = 35) (No Doppler)
1	Crew of 1-5	100%	100%	87%
	Crew of 2-5,3 or 5,5	94%	100%	82%
2	Crew of 1-5	100%	100%	94%
	Crew of 2-5,3 or 5,5	100%	100%	86%
3	Crew of 1-5	100%	100%	92%
	Crew of 2-5,5	11%	29%	10%
4	Crew of 1-5	100%	100%	88%
	Crew of 2-NA	-	-	-
5	Crew of 1-5	100%	100%	94%
	Crew of 2-5,3	60%	33%	59%
6	Crew of 1-5	100%	100%	88%
	Crew of 2-5,3	60%	33%	66%
Mean Percentage of Correct Estimates	Crew of 1	100%	100%	91%
	Crew of 2	68%	59%	63%

[1]Accuracy expressed as the percentage of estimated skill levels in agreement with criteria skill levels.

TABLE 15. ESTIMATES OF CHANGES TO 328XX CAREER FIELD REQUIRED TO SUPPORT THE PROPOSED DESIGN

	Changes to 328XX		Types of Changes			
	No	Yes	New AFSC	Combine AFSCs	New Shredout of AFSC	Expand Present AFSC
Group A (n=17) (APN-147)	76% (n=13)	24% (n=4)	1	2	1	
Group D (n=18) (Other Doppler)	88% (n=7)	12% (n=1)				1
Group X (n=35) (No Doppler)	83% (n=29)	17% (n=6)			3	3

APN-147 was introduced to the operational inventory, it has been maintained by technicians with AFSC designations of 301X1, 301X4, and, since 1972, by technicians with the 328X4 designation (avionic inertial and radar navigation system technician). The question on career field was defined in present career field designators (328X4) to avoid the potential confusion of asking technicians to recall AFSCs they may never have been acquainted with. This would be especially true for technicians with fewer than five years of Air Force experience. The responses to this item reflect accurate responses in terms of current AFSCs. Table 16 presents the responses to the more detailed question of which AFSC would be required to maintain the proposed equipment. The responses were between 94 percent and 100 percent correct over all groups.

TABLE 16. ESTIMATES OF AIR FORCE SPECIALTY CODE REQUIRED TO SUPPORT THE PROPOSED DESIGN

	328X0	328X1	328X2	328X3	328X4*	Other
Group A (n=17) (APN-147)	-	6% (n=1)	-	-	94% (n=16)	
Group D (n=8) (Other Doppler)	-	-	-	-	100% (n=8)	
Group X (n=35)	-	3% (n=1)	-	-	97% (n=34)	

*Actual AFSC performing maintenance on APN-147.

Task Difficulty Estimates

The difficulty of each defined maintenance task was rated by the technicians in comparison to the degree of difficulty they had experienced with other avionics maintenance tasks. These ratings were compared to the criteria ratings and accuracy values were calculated. The accuracy values are contained in Table 17. In almost all cases the accuracy values for the groups were less than 1.00 indicating that the technicians generally rated the maintenance tasks for the doppler signal processor as being less difficult than the difficulty ratings assigned by the criteria data judges. A Kruskal-Wallis one-way analysis of variance revealed that there were no significant differences in accuracy among the three groups ($H = 1.51$, $p<.50$).

Design Evaluations

The evaluations of the design presented in the Engineering Description Package were not productive. There was very little agreement between the design features identified by the estimators and those features identified by the manpower and maintenance criteria judges as features which would adversely affect maintenance time, safety, errors, and difficulty. Overall, 68 percent of the estimators did not identify any design features which would adversely affect maintenance. Since the manpower and maintenance judges did identify design features which did have an adverse affect on maintenance time, safety, errors, and difficulty, the estimators' "no" responses represent inaccurate responses. Of the 32 percent of estimators responding, only a small percentage of their responses (less than 3 percent) agreed with the criteria. It appears that maintenance technicians using early stage engineering data cannot identify features of the proposed equipment which would adversely affect maintenance time, safety, errors, and difficulty.

Training Estimates

Training Time. The focus of analysis for the training estimates was the accuracy of training time estimates for various training topics and types of training (see Tables 4 and 5). Maintenance technicians in groups A, D, and X as well as avionics instructors in Group K overestimated the training times for all types of training. An examination of the accuracy values for training time estimates contained in Tables 18 and 19 reveals the magnitude of the overestimates. The accuracy values for training time estimates for the general training topics (Table 18) represent overestimates ranging from 71 percent (accuracy value = 1.71) to 887 percent (accuracy value = 9.87). The accuracy values for estimates of training time for maintenance training topics (Table 19) represent overestimates ranging from 148 percent (accuracy value = 2.48) to 764 percent (accuracy value = 8.64). Although no further analyses were conducted because of the large degree

TABLE 17. ACCURACY OF TASK DIFFICULTY ESTIMATES

Systems Experience Groups	Accuracy Values*						Group Mean Accuracy
	Task 1	Task 2	Task 3	Task 4	Task 5	Task 6	
Group A (n = 17) (APN-147)	.53	.66	.97	.64	.56	.72	.68
Group D (n = 8) (Other Doppler)	.69	.61	.98	.83	.58	.85	.76
Group X (n = 35) (No Doppler)	.69	.57	1.03	.81	.65	.89	.77
Mean Accuracy Values	.64	.60	1.01	.76	.62	.84	

$$\text{*Accuracy Value} = \frac{\text{Estimated Difficulty Rating}}{\text{Criteria Difficulty Rating}}$$

TABLE 18. ACCURACY VALUES[1] FOR ESTIMATES OF GENERAL TRAINING TOPICS

	Accuracy Values for Training Time Estimates		
	Resident	OJT	FTD
Group A (n=14) (APN-147)	2.60	2.84	6.26
Group D (n=6) (Other Doppler)	2.52	1.71	6.69
Group X (n=20) (No Doppler)	9.87	2.52	7.26
Group K (n=20) (Avionics Instructors)	1.85	3.34	5.85

[1]Accuracy Value = $\frac{\text{Estimated Training Time}}{\text{Actual Training Time}}$

TABLE 19. ACCURACY VALUES[1] FOR ESTIMATES OF MAINTENANCE TRAINING TOPICS

	Accuracy Values for Maintenance Training Time Estimates		
	Resident	OJT	FTD
Group A (n=14) (APN-147)	4.71	4.16	3.37
Group D (n=6) (Other Doppler)	7.70	5.55	8.64
Group X (n=20) (No Doppler)	5.26	4.83	4.43
Group K (n=20) (Avionics Instructors)	2.48	5.95	2.67

[1]Accuracy Value = $\frac{\text{Estimated Training Time}}{\text{Actual Training Time}}$

of inaccuracy for the training time estimates, it does appear that Group K, avionics instructors, produced resident training time estimates which were relatively more accurate than the resident training time estimates produced by the other groups. This does not appear to be the case for OJT and FTD training time estimates.

Training Facilities/Equipment. Estimates of training facilities/ equipment were included in the study to determine if estimators could identify unusual training facilities or equipment to support the proposed equipment design. The training criteria judges, however, indicated that no unusual training facilities or equipment was necessary to support the proposed equipment. They identified the types of training equipment necessary to support most types of avionics training such as mock-ups, hot mock-ups, schematics, and block diagrams. The estimators also identified the usual types of avionics training equipment and, as a result, their responses agreed with those of the training criteria judges. However, this information provided no evidence as to the capability of expert estimators to determine unique training facility/ equipment requirements based upon early engineering data.

Engineering Description Package Evaluation

The overall usefulness of the Engineering Description Package was rated on a 100-point scale ranging from 0, "Not Very Useful," through 50, "Useful," to 100, "Very Useful." The mean rating given by the technicians was 35.8, somewhere between "Not Very Useful" and "Useful." When asked to rate the amount of detail in the package, 47 percent felt there was too little detail, 39 percent felt there was sufficient detail, 7 percent felt too much detail was presented and 7 percent did not respond. When asked if the diagrams and illustrations were helpful, 76 percent responded yes, 15 percent responded no, and 8 percent had no response. When asked about the number of illustrations and diagrams to be included in future summaries, 54 percent felt the number should be increased, 3 percent felt the number should be decreased, 31 percent felt the number should be about the same, and 12 percent had no response. With respect to length, 8 percent of the technicians felt the summary was too long, 36 percent felt it was too short, 47 percent felt it was about right, and 8 percent did not respond.

An improved engineering description package for this study would include more detailed information (if available in the early stages of equipment design) and more information on the proposed location of the equipment on the aircraft. The package would also contain additional information on test, alignment and adjustment procedures. Based on the technicians' comments the improved package would have a shortened introduction as well as less information on the theory of operation. The use of graphic means of presenting information would be retained and the number of graphics, such as diagrams and illustrations, would probably

be increased. With these changes incorporated, the improved engineering description package would be approximately the same length as or slightly longer than the original package.

DISCUSSION

The results of this study demonstrate that avionics technicians using an expert estimation technique can make accurate and reliable estimates of maintenance man-hours, crew sizes, crew skill levels and required AFSCs for proposed avionics equipment. These results provide additional validity data regarding the use an expert estimation technique for collecting these types of manpower data during the early stages of system design.

More detailed analyses of the man-hour estimates were made to determine if the qualifications of the technicians or the quantities of technicians affected the accuracy of the estimates. Statistically there were no significant differences in accuracy among the three systems experience groups. This finding should not be too surprising since these technicians, although varying in systems experience, represented closely related AFSCs within the avionics career field. Based on the results of this study, estimators from AFSCs or systems experience groups which are closely related to the proposed equipment or system will provide reasonably accurate estimates of maintenance manpower and man-hours. Further research should determine how much of a difference can exist between the estimators' systems experience and the proposed system before the accuracy of the estimated data deteriorates.

It also appears that it is not necessary to select technicians on the basis of length of systems experience. No significant differences in accuracy were found between those technicians with over four years of systems experience and technicians with four or fewer years of systems experience. These results may indicate that the experience base for making accurate estimates can be accumulated in less than four years.

The accuracy values for the man-hour estimates reported for the different sizes of groups were useful for determining the quantity of raters necessary to make accurate estimates. The mean accuracy values varied only slightly and with no clear pattern as the group size was increased. However, as group size was increased, the range of the accuracy values decreased indicating less variability in the estimates. This would indicate that the probability of obtaining extremely high or low man-hour estimates would be reduced for larger groups. For this reason, group sizes of at least 10 are recommended for use with the expert estimation technique.

Another consideration in recommending a particular quantity of raters is reliability. Interrater reliabilities increase as the number of raters increase. For the man-hour estimates, interrater reliability coefficients for individual estimators were .25 (Group D) and .27 (Groups A and X). Using the Spearman-Brown Prophecy formula with these coefficients the predicted interrater reliability coefficients for groups of 10 are .77 (Group D) and .79 (Groups A and X), respectively. For groups of 20 estimators, the interrater reliability coefficients are .87 (Group D) and .88 (Groups A and X), respectively.

In determining the number of estimators needed for a given project, the accuracy requirements for the project as well as available resources need to be considered. If a high degree of confidence in the estimates is needed and if resources are available, then it appears from the data that larger estimator groups should produce estimates with less variability and higher interrater reliability. In no case, however, should fewer than 10 estimators be used to generate man-hour, manpower, and maintenance estimates.

The recommendations on qualifications and quantities of raters discussed above are based on the analysis of the man-hour estimates. The criteria data for the man-hour estimates were the actual man-hours reported in maintenance records for the six maintenance tasks while the criteria data for the other items were ratings from the criteria judges. The man-hour criteria data were viewed as providing the most rigorous test of the expert estimate technique. The analysis of the man-hour estimates, therefore, became the basis of the recommendations for quantities and qualifications of raters. However, it is expected that the recommended qualifications and quantities of raters for man-hour estimates would also be used to collect the associated estimates of crew size, skill level, career field, and task difficulty.

The accuracy of the training time estimates was not encouraging. Generally, the technicians overestimated the amount of time required for the various types of training and training topics. One possible explanation is that the engineering description package did not contain enough detail to permit accurate training time estimates. Another possible explanation is that although the technicians were instructed to estimate the training times for the specific component, the doppler signal processor, they may not have been able to distinguish between the training time allotted to one component from the training time allotted to an entire system. The estimates may have reflected the longer training times associated with the entire system. Additional research seems in order regarding the impact on training time estimates of level of detail of engineering data and systems versus component level descriptions of proposed designs.

The questions regarding training facilities and equipment were included in the study under the assumption that the responses would produce estimates of unusually complex or expensive training equipment as well as more routine training aids. However, estimates of required training facilities/equipment could have applied equally well to the entire system or to any other avionics system. It is inconclusive at this time whether expert estimators using early engineering data can determine requirements for unique and complex training facilities or equipment.

It must be remembered that the results of this study were based on an avionics component and estimates from avionics technicians. This limits the application of the method to avionics technologies. Since avionics and electronics technologies are appearing in increasing numbers of Air Force systems this limitation may mean little in terms of the number of systems which could use the expert estimate method. In any case, as research on the expert estimate method includes additional technologies and technician career fields, it is probable that the method will be shown valid for a wider range of technologies.

A prototype guide for implementing the expert estimate method was developed and is available as AFHRL-TR-78-19 (Supplement 1), from the Defense Documentation Center and NTIS. The guide includes the five topics of: (1) developing the engineering description package of the proposed design; (2) developing the questionnaire to collect the estimate data; (3) selecting the technicians to serve as expert estimators; (4) collecting the estimate data; and (5) analyzing the data and preparing it for use by the engineer. The guide also includes a copy of the engineering description package and the data forms used by the expert estimators.

CONCLUSIONS AND RECOMMENDATIONS

The following conclusions and recommendations can be made based on the results of this investigation:

1. The expert estimate method employed in this study can be used to collect data on maintenance man-hours, crew size, skill level, Air Force Specialty Code, and task difficulty data for proposed equipment designs.

2. The estimates of training times and training facilities/ equipment were not accurate or productive. Further work is needed to develop techniques for estimating these training items during the early design phases of system development.

3. Technician estimators should be selected from the current systems most similar to the proposed system. The length of system experience does not seem to affect the accuracy of manpower and maintenance estimates provided the raters are at least qualified as 5-level maintenance technicians.

4. A minimum of 10 estimators should be used to reduce the variability in the estimates and to increase interrater reliability. The actual quantity selected would depend on the desired degree of confidence and available time and financial resources.

5. Refinements were made to the expert estimation technique developed by Whalen and Askren (1974). A prototype user's guide describing the refined method was developed and is available as AFHRL-TR-78-19 (Supplement 1), from the Defense Documentation Center and NTIS.

6. Additional research is needed using an actual, rather than a simulated, engineering description, and entire system, rather than a component of a system, and an inclusion of technologies and AFSCs from areas other than avionics. This research would broaden the generalizability and increase confidence in the use of the expert estimate method.

REFERENCES

Air Force Manual 66-1, "Maintenance Management Policy." Washington, D.C.: Department of the Air Force, 1 November 1975.

Decisions and Designs, Inc. Handbook for decision analysis. NR-197-023. Advanced Research Projects Agency and Office of Naval Research, 1973.

Haines, D. B., & Gael, S. Estimating manning requirements for advanced systems: a survey of the defense industry. AMRL-TDR-63-110. Wright-Patterson AFB, OH. Behavioral Sciences Laboratory, 6570th Aerospace Medical Research Laboratories, November 1963.

Huber, G. P., & Delbecq, A. Guidelines for combining the judgments of individual members in decision conferences. Academy of Management Journal 1972, 15, 161-174.

Smith, R. L., Blanchard, R. E., & Westland, R. A. Acquisition of restore time data by subjective techniques. AMRL-TR-70-76, Wright-Patterson AFB, OH. Aerospace Medical Research Laboratory, March 1971.

Whalen, G., & Askren, W. B. Impact of design trade studies on system human resources. AFHRL-TR-74-89, AD-A009 639. Wright-Patterson AFB, Ohio. Advanced Systems Division, Air Force Human Resources Laboratory, December 1974.

☆U.S. GOVERNMENT PRINTING OFFICE: 1978- 771-122/51